The Human Penguin Project: Social Integration Protects Against Cold Climates

Hans IJzerman[a], Siegwart Lindenberg[b,c], İlker Dalğar[d], Sophia S. Weissgerber[e], Rodrigo C. Vergara[f], Athena H. Cairo[g], Marija V. Čolić[h], Pinar Dursun[i], Natalia Frankowska[j], Rhonda Hadi[k], Calvin J. Hall[g], Youngki Hong[l], Chuan-Peng Hu[m], Jennifer Joy-Gaba[g], Dušanka Lazarević[h], Ljiljana B. Lazarević[h], Michal Parzuchowski[j], Kyle G. Ratner[l], David Rothman[g], Samantha Sim[n], Cláudia Simão[o], Mengdi Song[m], Darko Stojilović[h], Johanna K. Blomster[p], Rodrigo Brito[q], Marie Hennecke[r], Francisco Jaume-Guazzini[s,t], Thomas W. Schubert[n,u], Astrid Schütz[v], Beate Seibt[p], Janis H. Zickfeld[p]

Affiliations:

[a]Université Grenoble Alpes, [b]Rijksuniversiteit Groningen, [c]Tilburg University, [d]Middle East Technical University, [e]Universität Kassel, [f]Universidad de Chile, [g]Virginia Commonwealth University, [h]University of Belgrade, [i]Afyon Kocatepe University, [j]SWPS University of Social Sciences and Humanities, [k]University of Oxford, [l]University of California, Santa Barbara, [m]Tsinghua University, [n]Singapore Management University, [o]ISPA Instituto Univeritário, Lisbon, [p]University of Oslo, [q]Lusófona University, [r]University of Zürich, [s]Pontificia Universidad Católica de Chile, [t]Universidad Mayor, Santiago de Chile, [u]Instituto Universitário de Lisboa, ISCTE-IUL, [v]Otto-Friedrichs-Universität Bamberg.

Corresponding Author: Hans IJzerman, Université Grenoble Alpes. Email: h.ijzerman@gmail.com. Code needed to run our analyses can be found here: https://osf.io/6yu5d/?view_only=3da753a58f22450aa82115b73f88db14. Data and materials are linked throughout the paper.

Abstract

Social ties in general, but having a wider variety of social ties (i.e., complex social integration) specifically, have been shown to create important physical and mental benefits and thus belong to highly important determinants of life chances. A novel, but hotly debated, approach to identifying drivers of complex social integration comes from social thermoregulation theory. The theory is derived from homeothermic animals other than humans and pertains to the idea that modern human relationships are pleisiomorphically organized around body temperature regulation. In two studies (a pilot study and a study covering twelve countries) we first identify major drivers of core body temperature through powerful exploratory methods. After identifying complex social integration as one of the key predictors of core body temperature, we identify a path model through a split-half cross-validation method that shows that colder climates relate to higher levels of complex social integration, while such complex social integration in turn relates to higher core body temperatures. We infer that – despite modern conveniences like clothing and heating – people still rely on social warmth to buffer their bodies against the cold. Our studies do not only contribute to a deeper understanding of social network formation, but also provide direction for how social relations contribute to health and well-being.

Keywords: *Social integration, Social Thermoregulation Theory, Attachment Theory, Embodiment, Machine Learning*

The Human Penguin Project: Social Integration Protects Against Cold Climates

One key motivating force for bonding across mammals is their need to regulate body temperature. Without adequate temperature regulation, they die. Distributing body heat across conspecifics makes responding to environmental fluctuations in temperature less costly energetically, with this distribution manifesting in a lowering of metabolic rate. Perhaps more surprisingly, since 2008, the literature on adult human beings has also shown the reliance of feelings of trust (or psychological warmth) on temperature regulation (for an overview, see IJzerman & Hogerzeil, 2017). But this literature has not been without controversy. First, some effects have and some effects have not replicated. Furthermore, some uncertainty may still exist as to how the sharing of body heat extends to the fuller spectrum of modern human relationships. Third and finally, it is unclear whether "social warmth" indeed protects people's core body temperatures from the cold.

In this report, we sought to resolve these outstanding issues through a high-powered pilot study and an even higher-powered cross-national study. We first explore *which* variables are most potent in predicting core temperature through a powerful exploratory method we borrow from artificial intelligence called supervised machine learning. In our main study, we then again used the same exploratory method, after which we conduct a split-half cross-validation of a path model to know *how* the earlier identified variables relate to core temperature through a training and testing dataset (ensuring the robustness of our results). Altogether, convergence from our machine learning results and our path models support a strong relationship between people's environmental temperature (operationalized as distance from the equator), their levels of social integration across different relationships, and their core body temperatures.

Social Thermoregulation as Key Facet of Human Social Attachments

Having high quality social relationships is one of the biggest predictors of one's health (Holt-Lunstad et al., 2010). Although scholars dating back to Hippocrates have understood that disturbances in health closely relate to dysregulated body temperature (Benzinger, 1969; Minard et al., 1964), evidence for the link between thermophysiology and social relationships in humans has only begun to accumulate since 2008. In its most elementary form homeotherms (=warm-blooded animals) maintain temperature at homeostatic levels in various ways, mounting numerous defense systems, like yawning, panting, or sweating when temperatures increase or shivering when temperatures decrease (Gallup & Gallup, 2008; Jansky, 1973). Temperature regulation usually extends beyond internal regulation by turning to others. The reliance on conspecifics has been remarkably asymmetrical: Temperature increases are typically resolved by the organism itself (e.g., through internal regulation like yawning or through behavioral thermoregulation like getting into colder water) because increases can be immediately threatening for survival. In contrast, because temperature decreases are not immediately dangerous, regulation back to homeostasis is often "outsourced" to conspecifics through huddling.

Strong evidence for these ideas can be found in studies across homoeothermic (nonhuman) animals. In rodents, social thermoregulation has been shown to be one of the most important motivating forces behind group living, especially when temperatures drop (Ebensperger, 2001). The Octodon Degus (a Chilean rodent) for example used 40% less energy and achieved a higher surface temperature when housed with three or five others (vs. alone; Nuñez-Villegas et al., 2014). Studies of vervet monkeys show somewhat more complex mechanisms, with larger social networks buffering their core temperatures from the cold (McFarland, Fuller et al., 2015), while even grooming a dead vervet monkey's pelt insulates its body against temperature variations (McFarland, Henzi et al., 2015).

In humans, *social thermoregulation* extends beyond huddling. More specifically, social thermoregulation theory explicates how what English speakers intuitively know as "social warmth", that is, trustworthiness and social predictability, relies on and growths out of more ancient needs of physical warmth. The outgrowth has likely happened to avoid redundancy: As in earlier days, some aspects (like regulating body temperature) were crucial for survival, it was efficient to "reuse" those brain areas (for thermoregulation) for other purposes (for social interaction; Anderson, 2010; Satinoff, 1982, 1983). The connection between social behavior and thermoregulation is likely not incidental: Like other homeotherms, older day humans simply needed to stay physically proximate to stay warm.[1]

1 Past research (for example by IJzerman & Semin, 2009) has attributed thermoregulatory effects to metaphor theories (Lakoff & Johnson, 1983, 1999). We now believe this is wrong. Metaphor theories posit that connections between concrete experience and abstract concepts occur because activation in one area (e.g., for social interaction) becomes associated with superficially related areas (e.g., for physical warmth) in the brain. Metaphor theories also suggest unidirectionality (e.g., manipulating warmth should lead to closeness, but manipulating closeness should not lead to warmth) and cross-cultural universality (the metaphor should hold the world over). It is now clear however that the central predictions of metaphor theories (bi-directionality and cross-cultural universality) have been falsified. Manipulations of loneliness do lead to changes in temperature perceptions (e.g., IJzerman & Semin, 2010) while linguistic metaphors related to warmth and affection are not universal, with many languages around the equator not showing the WARMTH is AFFECTION metaphor (Koptjesvkaja-Tamm, 2015). Furthermore, while some effects appear to support metaphor theories by showing an overlap of insular cortex activation for social and physical warmth (e.g., Inagaki & Eisenberger, 2013), other research has clarified that it is *subjective* temperature changes that lead to insular cortex activation (Craig et al., 2000) while many social and thermoregulatory effects already happen at much lower levels, like at the level of the medial pre-optic area of the hypothalamus (Boulant, 2000). The more appropriate way to conceptualize the underlying organization of neural systems related to social thermoregulation is as a "hierarchical prediction machine", where higher order areas (e.g., those related to insular cortex activation and monitoring social contact) help foster more efficient activity at lower levels (e.g., regulating temperature and social contact at the hypothalamic level; Clark, 2013; Friston, 2008).

From that perspective, it may not come as a surprise that the aggregate evidence has come in favor of this evolved relation between social and physical warmth (IJzerman & Hogerzeil, 2016; IJzerman, Janssen, et al., 2015; Schilder et al., 2009; Williams & Bargh, 2008). Evolutionary pressures that used to be related to infant survival (e.g., feeling cold and wanting to be held) likely form the basis for an evolved template for mental (attachment-like) models concerning the relationship between physical and social warmth. This relationship is confirmed in research showing that an individual's needs to socially thermoregulate correlate highly (and negatively) with attachment avoidance (Vergara et al., 2017). Furthermore, the priming of (lack of) trust is as asymmetrical as the underlying physiological systems: Priming trust leads to higher temperature perceptions when temperatures are low, but not when temperatures are high (Ebersole et al. 2016; IJzerman et al., 2016).

This relationship can also be observed in its most basic form in the so-called "Strange Situation", in which infant researchers observe the behavior of an infant to separation with the mother. When the mother leaves the room in this ethological observation, infants' skin temperature drops, and peripheral temperature only returns to baseline once the mother returns (Mizukami et al., 1990). Similar effects can be observed in adults: Students' peripheral temperatures drop when they feel socially excluded (IJzerman et al., 2012).

Nevertheless, some controversy has arisen around the topic, as not all effects in this literature replicate.[2] To help solve these issues, we generate our research questions from the basic

2 Many published studies are heavily "underpowered" (i.e., samples too small to be able to support the tested hypothesis) or fail to replicate (Open Science Collaboration, 2015). It may then also not come as a surprise that also in the field of thermoregulation some effects failed to replicate (Vess, 2012; LeBel & Campbell, 2013). Yet, other effects did replicate (Schilder et al., 2014; Ebersole et al., 2016; IJzerman & Semin, 2009; IJzerman et al., 2016; Inagaki, Irwin, & Eisenberger, 2015), while some effects have been obtained with considerably samples, like those between 100 and 500 (e.g., IJzerman, Janssen, et al., 2015; Van Acker et al., 2016), or even around 30,000 (Hong & Sun, 2012) and above 6 million (Zwebner et al., 2013). Finally, considerable converging evidence exists on other species and on human biological functioning that attests to the importance of social thermoregulation (for reviews,

assumption that the availability of modern conveniences to regulate temperature has been so brief that the evolved link between physical on social warmth is likely to *still* lead to strategies that help buffer individuals from the cold through their social networks - even via relationships that do not typically permit physical touch.

From research, we already know that feeling cold increases the need to socially connect (Van Acker et al., 2016). There is only one pilot study demonstrating a relationship between network quality and higher core temperatures in humans, showing that having a greater quality social connection is positively correlated to core temperature (Inagaki et al., 2016). But on the basis of the existing literature, it is not at all clear *whether* social connections protect against the cold, *which types* of social contacts protects against the cold, *whether* social contacts are more prominent in predicting core temperatures than other known predictors. We therefore first explore which variables are crucial in predicting core temperatures. For this, the theoretical background and the research questions derived from it function mainly as a basis for including the variables in our study.

Choosing Prediction over Explanation

To achieve our goals, we relied on an exploratory approach first, a method unusual for psychological research which has typically focused on testing theoretically derived hypotheses. Several researchers have now argued that our focus on identifying complex prediction-focused models together with flexibility in data analyses have led to what is called "overfitting" (mistaking noise for a real signal by fitting an overly complex model to existing data). This

see IJzerman & Hogerzeil, 2016; IJzerman, Coan et al., 2015; Terrien et al., 2011). There are, of course, a number of reasons why effects may not replicate. The first is that the theory is incorrect, and results rely entirely on false positives, caused by a combination of small sample studies and data-contingent analyses. The second is that the replication was not well executed. We suspect that neither is the case here. Instead, we think that the problems lie in still insufficiently specified models behind the theory, and thus with specific manipulations and specific measurements in specific contexts (with small samples).

process of overfitting is arguably one of the lead causes of the replicability crisis (for an overview, see Yarkoni & Westfall, in press). One way to reduce the problem of overfitting is by relying on exploratory analyses before moving on to more mechanism-based explanations of human behavior (Brandt, 2017; IJzerman, Pollet et al., 2017; Yarkoni & Westfall, in press).

Researchers in artificial intelligence have advocated an exploratory analysis technique that deduces patterns from data called *(supervised) machine learning.* A very popular and widely tested approach to supervised machine learning is a method called “random forest” (Breiman, 2001).[3] This method allows for measuring the relative contribution of each specific variable, considering all the variables used in a dataset to predict the outcome of a “signal” through a bootstrapping-type method, yielding a highly predictive accuracy. The outcome of this repeated sampling is captured in a variable importance list that indicates which variables are very likely to predict the outcome variable (for more technical discussions see Breiman, 2001; IJzerman, Pollet, et al., 2017; Jones & Lindner, 2015; Yarkoni & Westfall, in press).[4] This method thus has

3 Supervised machine learning differs from unsupervised machine learning in that the data patterns are derived by a “supervisory signal” (an outcome variable). In unsupervised machine learning, the algorithm infers a hidden function or pattern from the data, without regard to such a “signal” (which we typically refer to as dependent variable). The type of supervised machine learning we use relies on a regression function, very similar to the correlational analyses psychologists are accustomed to. However, the type of machine learning we used here 1) allows for non-linearity (which standard regressions cannot do without a priori specification), 2) does not presume direction (a positive or negative relationship), 3) has much less problems with collinearity, and 4) is agnostic which type of variable predicts the outcome, thus allowing the researcher to *classify before regressing* onto the signal (i.e., what psychologists typically refer to as dependent variable). Finally, the method we used is *conditional* random forest. Whereas typically, random forests are formed in a training dataset and then confirmed in a testing dataset, conditional random forests correct for error internally.

4As the name implies, the “forest” consists of many such “trees”. The method relies on “out of bag estimates” (bagging), which involves repeated sampling to form training datasets from an original dataset (Breiman, 1996; Bylander & Hanzlik, 1999). The rest of the datasets in each case (the test datasets) are used to evaluate the prediction power of the variable importance and trees trained on the training dataset. The forest is then aggregated with each tree getting a “vote”, which constitutes a weight in the ensembled model that summarizes all information from the trees. Because the analyses are exploratory, it does not provide an effect size estimate but instead shows the relative importance of one variable over the other within the model, and as compared to random noise. The type of machine learning we used (*conditional* random forest) improves its predictive power throughout each iteration of the analyses.

the potential to reduce bias in analyses and is particularly useful when multiple predictor variables are measured. Overall, it is very useful as an exploratory approach to identify variables for further confirmatory testing. For us that means that we will able to pinpoint which variables are most relevant in predicting core body temperature in Study 1 and 2 and will subsequently allow us to pinpoint a mechanism by testing a path model in a split-half cross-validation in Study 2's Phase II.

The Human Penguin Project Overview

We sought to accomplish our goals in two studies. We first ran an online pilot study (*N*=240) and a large, cross-national study (12 countries; *N*=1507) to identify which variables are most accurate in predicting people's core temperatures. We measured a number of known correlates of body temperature and a number of variables on social interaction that – based on prior research – should logically be related to core body temperature (e.g., nostalgia (Zhou et al., 2012) or attachment to homes (Van Acker et al., 2016))[5]. In selecting our variables, we were over- rather than underinclusive, as our first priority was to identify which variables are the most prominent predictors of core body temperature. In our second (main) study, we again first relied on supervised machine learning, after which we specified our path model through a split-half cross-validation method.

Method Pilot Study

5 We included a number of variables that have been found to relate to body temperature (stress; Marazziti et al., 1992; whether participants use medication) or those that have been known to relate to environmental temperature variations (nostalgia; Zhou et al., 2012; attachment to homes; Van Acker et al., 2016) or to metabolism and social network quality (like daily (diet) sugary drinks consumption; Henriksen et al., 2014). We were also overinclusive, asking about questions that relate to the regulation of stress (and could thus relate to body temperature) like self-control (Tangney et al., 2004), attachment (Fraley et al., 2000), and access to one's own feelings and bodily states (alexithymia; Kooiman et al., 2002). Finally and most importantly, we included measures to assess the quality of people's social networks ("networksize", "socialembedded", and a measure on complex social integration, "CSI"; Cohen et al., 1997).

Participants and Procedure. Our questionnaire was programmed into the online platform Qualtrics and we collected data from mTurk (N = 143) and Prolific Academic (N = 148). Participants were requested to complete the survey between 9-11am, not to eat or drink anything warm or cold for 10 minutes preceding the survey, and not to have exercised an hour preceding the survey. Because the sample was relatively small, we excluded all those participants that did not adhere to these guidelines (mTurk $N = 3$; PA $N = 48$). The total remaining N for the pilot study was 240.

Participants entered the survey, where they were requested to fill in a number of different questionnaires. At the beginning of the questionnaire (Measurement 1) and at the end of the questionnaire (Measurement 2), they were requested to measure their own oral temperature with an oral thermometer, take a picture of the thermometer (with date, time, and Measurement– 1 or 2 – included; for an example photo uploaded by our participants, see Figure 1).

Survey details. Our dataset included a number of scales relevant to thermoregulation. In order to assess the importance of complex social integration and distance from the equator, we measured known correlates of core body temperature or behavior in response to temperature fluctuations, like self-reported stress ("stress" in the forest plot; Cohen & Wills, 1985), nostalgia ("nostalgia"; Routledge et al., 2008), attachment to homes ("attachhome"; Harris et al., 1996), daily sugary drinks consumption ("gluctot"; Henriksen et al., 2014) and diet drinks consumption ("artgluctot"; Henriksen et al., 2014), known benchmarks of core body temperature, like sex ("sex"), height ("height"), weight ("weightkg"), and whether they used medication ("meds"; Hills & Rahimtulla, 1965; Peters, 1986).

We also included variables that potentially influence core body temperature and quality of the social network in other ways, like feelings of agency, measured through self-control ("selfcontrol"; Tangney et al., 2004), attachment ("avoidance" and "anxiety"; Fraley et al., 2000), and access to one's own feelings and bodily states (alexithymia subscales "EOT" and "DIDF"; Kooiman et al., 2002). The complete scales, reliabilities, and averages per scale per site can be accessed on our project page. Finally, we looked up the minimum temperature ("mintemp") and average humidity of the day ("avghumidity") participants completed the survey based on their IP address by using a weather history site (http://www.wunderground.com/history/), which bases weather on the nearest airport. At the end of the survey, participants were thanked and debriefed for their participation.

Finally and importantly, we included measures on people's social networks ("networksize", "socialembedded", and a measure on complex social integration, "CSI"; Cohen et al., 1997). CSI includes an inventory of the following ties: Relationships with spouse, parents, parents-in-law, children, other close family members, neighbors, friends, workmates, schoolmates, fellow volunteers (e.g., charity or community work), members of groups without religious affiliations (e.g., social, recreational, professional), and members of religious groups. One point was assigned for participation in each kind of relationship for which respondents reported that they spoke (in person or on the phone) to someone in that relationship at least once every 2 weeks.

Analyses and Results Pilot Study

Our method, conditional random forest, consists of an algorithm that classifies data points by weighting "votes" of predictions of a potential hypothesis underlying the relationship between variables. It then creates “decision trees” that indicate which variables get the most weight in relevance for the “signal”. The order in which these decisions are taken are represented by the “levels” of the tree. The path from the root of the tree to a node is a series of decisions and the node is then tagged by the prediction power of such a path. Given enough data (and enough predictors) these decision trees are very flexible, as the algorithm explores all possible relationship between predictor variables and signal that could be generated from the data. Each of these decision trees then allow for specification of the strength of predictor of variables based on a “desired outcome value” (the “supervisory signal”).

As the name implies, the “forest” consists of many such “trees”. The method relies on “out of bag estimates” (bagging) and conditional random forests immediately engage in error-correction during the process, which means that, as a result, no training vs. test dataset is needed as happens in typical random forests. The forest is then aggregated with each tree getting a “vote”, which constitutes a weight in the ensembled model that summarizes all information from the trees. This then created a *permutation variable importance* (i.e., the list with relative importance of each variable in predicting the outcome variable).

For creating the classification trees, we relied on R packages tree (Ripley, 2016), lattice (Sarkar, 2017), plyr (Wickham, 2016), stargazer (Hlavac, 2015), and summarytools (Comtois, 2016). MTry is the numbers of variables (out of the total list of variables) sampled at each split. MTry is recommended to be the square root of the total number of predictors. For our pilot study, we ran the analyses twice, with an original analysis and a replication (with mtry = 7, trees = 1000, for seeds 666 (original) and 667 (replication); link to script: https://osf.io/djh4b/ and to data: https://osf.io/smr3g/). The chance for overfitting is further reduced is by examining the stability of the two analyses through a Spearman Rank correlation between the forests. In this case, the model was very stable, because the Spearman Rank $r = .96$ (for more details about the procedure, see IJzerman, Pollet et al., 2017; for link to variable orders, see https://osf.io/2qvhb/ and for output, see https://osf.io/5cs5n/; see Figure 2 for one of the two dotplots).

The outcome of our conditional random forest analyses in our pilot study was that core temperature is best predicted by (in order of importance) height > CSI > weight > sex (see Figure 2). These variables exceed the "random noise threshold" in our forest, suggesting that these (and not others) differ from random noise in our dataset (as indicated by them not exceeding the red line that defines what differs from random noise in a dataset; Strobl et al., 2009).

Discussion Study 1

Our first study showed that the best predictors of core body temperature in our online samples of Prolific Academic and mTurk were height, CSI (complex social integration), weight, and sex (in that order). Beyond known benchmarks like height, weight, and sex, we discovered that CSI was one of the most important predictors of core body temperature (with CSI positively relating to core body temperature). Why could this be so? There may be different reasons, but the risk factor of a low variety of relationships may be paramount. In our complex social world, rejection and loneliness are common experiences because of the tendency to compare ourselves to others and because of the fragility of many relationships in life - from work to recreation to home to friendships. As a result, putting all eggs into one basket, that is, staking too much in any particular type of relationship is risky and potentially isolating if conflict arises or we do not feel we are meeting the standards in a particular life domain (Crocker et al., 2003; Deci & Ryan, 1995; Kernis, 2003; Steverink & Lindenberg, 2008).

Given the unknown stability of any particular domain of life, past research has found that having a wide-range of strong social ties – or, higher levels of complex social integration - is particular important for health and well-being (Cohen et al., 1997; Seeman, 1996). This is why some of the strongest evidence for the buffering effects of a web of social ties (versus putting too much weight in the strength of particular social tie) comes from studies that have assessed levels of CSI. This measure asks questions related to whether people are in regular contact with people in multiple facets of their lives (parents, relatives, close friends, colleagues, and so forth).

It is likely that the mechanisms are in place to achieve higher core body temperature through diverse social contacts. Recall that we already know that feeling cold increases the need to socially connect (even via email and phone; Van Acker et al., 2016) and that one pilot study has shown a relationship between feeling more connected socially and higher core temperatures in humans (Inagaki et al., 2016). Furthermore, we know that people project their relationships even onto inanimate objects like consumer products (Hadi et al., 2012; IJzerman Janssen, et al., 2015; Rotman et al., 2016). It is also known that neural areas related to thermoregulation overlap considerably with those related to (even more complex) social behaviors (Satinoff, 1982, 1983). Given that having higher levels of CSI is the most established buffer against loneliness, and previous work suggests that loneliness is related to physical body temperature, it makes sense that CSI would be positively associated with body temperature.

In our second, larger and cross-national project, we again sought to use the same exploratory method and then proceed with a split-half cross-validation analysis that could help us specify a path model. We now expect that CSI will again turn up as an important predictor of core body temperature, but we also sought to explore whether CSI protects against the cold.

The Human Penguin Main (Cross-National) Project

Research Summary

In our pilot study we found that CSI was one of the most potent predictors of core body temperature. In our cross-national project we sought to investigate whether CSI protects against the cold. We did so by again relying on supervised machine learning to identify predictors for CSI and again for core body temperature. We suspected that people who live in colder environments (= further from the equator) would need to rely more on social contacts to keep warm. Because the links to date have been underspecified, we again used powerful exploratory analyses to specify the exact relationships. We did so both for CSI (an important predictor of CBT in our pilot study) and for core body temperature, before we moved on to our path models.

Participants

In our cross-national study we tested the interrelationship between climate, complex social integration, and core body temperature on a fairly large scale, including 12 different countries on 3 different continents, and 1,507 participants. We report all our exclusions in our data handling section. We also report all of the variables we measured in our study, except for two questionnaires that researched to develop and validate those scales. The first author recruited collection sites through personal contacts and through “the ManyLab” (https://osf.io/89vqh/). Participants again completed a variety of online questionnaires at home or in the lab (depending on site). Answering the questionnaire took approximately 35 minutes in total.

Method Cross-National Study

Samples. We collected data via University of Oxford (UK; $N = 137$, 56.2% female, $M_{\text{birthyear}} = 1985.43$; $SD_{\text{birthyear}} = 13.51$), University of Belgrade (Serbia; $N = 164$, 80.5% female,

$M_{birthyear}$ = 1993.73; $SD_{birthyear}$ = 4.91), Singapore Management University (N = 135, 56.2% female, $M_{birthyear}$ = 1993.80; $SD_{birthyear}$ = 1.54), Tsinghua University (China; N = 174, 62.2% female, $M_{birthyear}$ = 1993.68; $SD_{birthyear}$ = 6.41), University of Zürich (Switzerland; N = 37, 72.5% female, $M_{birthyear}$ = 1987.57; $SD_{birthyear}$ = 8.72), Virginia Commmonwealth University (United States; N = 150, 78.8% female, $M_{birthyear}$ = 1992.62; $SD_{birthyear}$ = 4.70), University of Kassel (Germany; N = 105, 69.8% female, $M_{birthyear}$ = 1990.31; $SD_{birthyear}$ = 7.82), University of California, Santa Barbara (United States; N = 108, 63.8% female, $M_{birthyear}$ = 1995.82; $SD_{birthyear}$ = 1.71), University of Lusófona (Portugal; N = 18, 33.3% female, $M_{birthyear}$ = 1984.12; $SD_{birthyear}$ = 11.92), University of Chile (Chile; N = 34, 62.9% female, $M_{birthyear}$ = 1979.33; $SD_{birthyear}$ = 13.16), University of Southampton (United Kingdom; N = 6, 50.0% female, $M_{birthyear}$ = 1992.17; $SD_{birthyear}$ = 1.60), Otto-Friedrichs-Universität Bamberg (Germany; N = 40, 69.0% female, $M_{birthyear}$ = 1982.11; $SD_{birthyear}$ = 14.67), Middle East Technical University (Turkey; N = 181, 65.7% female, $M_{birthyear}$ = 1992.42; $SD_{birthyear}$ = 5.00), University of Oslo (Norway; N = 85, 69.4% female, $M_{birthyear}$ = 1992.31; $SD_{birthyear}$ = 6.42), and SWPS University of Social Sciences and Humanities (Poland; N = 133, 86.6% female, $M_{birthyear}$ = 1986.18; $SD_{birthyear}$ = 8.89). Our total sample (after data exclusions) consisted of N = 1507 (68.9% female, $M_{birthyear}$ = 1990.95; $SD_{birthyear}$ = 8.45).

Procedure. We created one central survey, which was translated and back translated keeping in mind loyalty to the original meaning. All surveys were programmed into the online survey platform Qualtrics. Participants were run online or in the lab across our different sites. Participants were again requested to complete the survey between 9-11am in their local time zone, not to eat or drink anything warm or cold for 10 minutes preceding the survey, and not to have exercised an hour preceding the survey. To be sure, we again asked whether they did eat or drink anything warm or cold 10 minutes before the study ("eatdrink") or whether they had exercised an hour preceding the study ("exercise"). At the beginning and end of the task, participants again measure their own temperature with an oral thermometer of which they took a picture and uploaded this to our online platform (for an example, see Figure 2; descriptives, analysis script, and details of how the study was conducted at each site are available on our OSF project page; https://osf.io/mc5gu/).[6] In our main study, the range of CSI was 4-12 and the average was 7.41.

Survey details. We used the same questionnaires as our pilot study, but now added a few questions that may also bear relevance for CSI and core body temperature that pertain to the nature and structure of their relationships, like whether people are in a romantic relationship or not ("romantic"), how monogamous they perceive themselves to be ("monogamous"), and questions that pertain to the size of their online social networks ("onlineid" and "attachphone"), while we also recorded participants' longitude and latitude via a standard option available in Qualtrics ("longitude"; we calculated latitude into equator distance "DEQ"). Finally, as the number of social contexts in which people are socially engaged may differ widely between cultures and language coding for "warm" and "cold" (Koptjevskaja-Tamm, 2015), we also included proxies for cultural influences with dummies for "language family" (Indo-European, Sino-Tibetan, and Uralic). Because cultural influences may be similarly large within a language family when the same language is spoken in highly different longitudinal locations (such as English spoken in the US versus that spoken in Singapore), we also included degrees longitude.

6 There was one exception to the usage of the oral thermometer: Participants at UCSB used a temporal artery thermometer. To be sure, we ran the analyses also without participants from UCSB. The effects for the full mediation model remained the same: There was a mediation for participants with a relationship (95% CI [.0005 . 0015]), but not for participants without a relationship (95% CI [-.0001 .0004]), with a significant interaction between DEQ and having a romantic relationship or not onto CSI (B = -.02, t = -5.05, p < .01, 95% CI [-.03 -.01]). For analyses excluding UCSB sample, see https://osf.io/b6r9v/.

At the end of the survey, participants were thanked and debriefed for their participation. We again looked up minimum temperature of that day and average humidity of that day through their IP address and the weather history site.

Data Handling. Before analyzing the data, for each scale variable, we checked the questionnaire's reliability, and corrected labeling differences between sites where necessary (a complete file with all alterations can be requested from the first author). We then created a final "raw" dataset. Next, we reviewed all pictures that participants uploaded to our Qualtrics platform. We made 193 (mostly small) corrections to the CBT values, based on the picture participants uploaded. We also deleted 13 participants, as these participants uploaded either generic pictures or pictures that were irrelevant for our study. When no picture was uploaded, we kept the participant in our dataset. We also deleted participants from our dataset that reported core temperature values ("CBT" variable in the dataset) lower than 34.99 degrees Celsius, and participants that reported very unlikely core temperature values (e.g., 100 degrees Celsius). Our final sample consisted of 1507 participants. Because we had a far larger N than our pilot study, we were somewhat more liberal with our inclusion on the basis of time of day, and left participants in even when they were not within our requested time frame. Instead, we included the time of day at which they ended the survey ("endtime") as control in the random forest and then in our mediation analyses.

Analyses and Results Cross-National Study

Degrees of freedom or sample size may differ throughout due to missings in specific variables. We do not outline them here, but point to the data available from our project page. For our Cross-National Study, we split our analyses in two phases. In the first phase, we again relied on conditional random forests to specify variable importance, but now with both Complex Social Integration (CSI) and Core Body Temperature (CBT) as supervisory signals in two separate analyses. For both supervisory signals, we ran 8 versions (1 original and 7 replications; 4 different seeds, and 2 different levels of mtry), which ensured that we obtained the most stable model possible. The lowest Spearman Rank for CSI was r = .971 and for CBT was r = .912 (scripts, data, and results are all available on our project page: https://osf.io/mc5gu/).

The following were the most important predictors of CBT (Figure 3; in that order):

sex > endtime > langfamily > mintemp > equatordistance (DEQ) > CSI > longitude > heightm > meds[7]

The following were the most important predictors of CSI (Figure 4; in that order):

langfamily > longitude > romantic > equatordistance (DEQ) > CBT > mintemp > age

Our results thus unequivocally show that distance from the equator (DEQ) and complex social integration (CSI) are amongst the most important predictors of core body temperature (CBT), close to being as important as sex and “language family”, and more important than known benchmarks like height, weight, and stress. Furthermore, DEQ is also an important predictor of CSI (Figure 3). Because DEQ and mintemp correlate highly, we chose to retain DEQ in our further analyses (but comparable results are obtained when using mintemp).

Based on these initial results and based on the relevance of distance from the equator and social integration, we decided to test a mediation hypothesis through a split-half cross-validation

7 Note that both height and weight did not turn up as predictors of core body temperature in our cross-national project. In hindsight, this should not be surprising. Across homeotherms, within the same taxa, body size correlates with distance from the equator (something that has become known as “Bergmann’s rule” (Bergmann, 1847). Larger animals have a lower surface to body ratio, making them better able to stay warm in colder climates (something that is also true for modern humans; Foster & Collard, 2013). We found a comparable correlation for either height (*r* (training set) = .116, $p = .002$; *r* (testing set) = .192, $p < .001$) or weight (*r* (training set) = .118, $p = .002$; *r* (testing set) = ..180, $p < .001$) with DEQ in our sample.

method, asking whether people further from the equator have lower core temperatures first. We created a training and test dataset through a random number generator in SPSS. We tested a model in our training dataset, which we then sought to confirm in our testing dataset.

For our mediation model, we based ourselves on our machine learning results. We conducted more traditional regression analyses to further understand the relationship between our most important variables. We did so again in the most robust way possible, by creating a training dataset (https://osf.io/t25d5/) and a test dataset (https://osf.io/wa8xk/), and examining which hypotheses survived analyses in both datasets. To provide the highest informational value possible, we report here the analyses over the entire dataset (https://osf.io/txskz/; for details about each step of the analyses, see our project pages).

In all analyses that follow, we control for variables in both steps of the mediation that showed to be important in our conditional random forest, but which were not central to our hypotheses (participants' language family, sex, and the time of day they finished the study). Nevertheless, all analyses showed the same pattern *without* these controls.[8] In a regression, DEQ shows a robust relation with CSI (with $B = .015$, $t = 7.03$, $p < .001$, 95% CI [0.011, 0.019]).[9] In turn, CSI is a positive predictor of core body temperature (CBT; in a regression; $B = .043$, $t = 4.32$, $p < .001$, 95% CI [0.023, 0.062]), while DEQ is a negative predictor of CBT ($B = -.0045$, $t = -5.61$, $p < .001$, 95CI [-0.006, -0.0029]). These results thus show a robust relationship between DEQ, CSI, and CBT : Having a more varied active set of social relations increases core body temperature slightly, while distance from the equator decreases it. People further away from the

---

8 For analyses with controls, see https://osf.io/97t39/ and without controls see https://osf.io/4q4sg/.

9 Because of missing values, degrees of freedom differ per analysis. The exact degrees of freedom can be obtained from our analyses output on our OSF project page (https://osf.io/2rm5b/).

equator have a more varied set of social relations..[10,11]

Because our data is cross-sectional and thus only allows indirect causal statements, we conducted a number of extra analyses to explore the interrelationship between DEQ, CSI, and CBT. These again followed the logic of relying on our training and then our test dataset. First of all, because having a romantic relationship was a key predictor of CSI, we explored the influence of romantic relationships and found that the effects differ for those who do and those who do not have a romantic relationship. There are different possibilities for the mechanisms involved. Having a romantic relationship could provide the individual with an initial safe haven, making her less inhibited to explore and connect closely to others in various social contexts that can help protect core temperature, as one would predict on the basis of attachment theory (Bowlby, 1969). By contrast, having a romantic relationship could be a proxy for a *reduced* urgency to derive social warmth from CSI, in which case DEQ would be a weaker predictor of CSI for those with a romantic relationship.

We see the former conjecture (i.e., explore and connect) supported in the training, testing, and overall data (see left [romantic relationship] and right [no romantic relationship] panels of

---

10We also explored whether network size had comparable effects, and, finally, we tested for robustness of our model by running both with and without the covariates that we identified on the basis of our machine learning models. We report only the 95% confidence intervals of additional mediation analyses here that were not reported in the main text. The exact analyses are reported on our project page. We ran the mediation also without controls and without "romantic relationship" as moderator, and the DEQ-CSI-CBT mediation was again significant (CI95 .0003 . 0010). When we included our controls sex, time of day, and language family (but without moderation by relationship status) the DEQ-CSI-CBT mediation remained significant (CI 95 .0004 .0010). When we moderated for relationship status (but without the specified controls) the DEQ-CSI-CBT mediation only worked as reported in the main text for those with a romantic relationship (CI95 .0009 .0020). In this case, the DEQ-CSI-CBT mediation (without controlling for the specified variables) seemed partially reversed for those without a relationship (CI95 -.0012 -.0002). More specifically, the reversal of this mediation meant that there was a negative relationship between CSI and CBT for those without a romantic relationship (B = -.0083, t = -2.40, p = .02). Given the sample size, the p-value of .02, and the fact that this was the only situation in which this reversal occurred, we consider this a fluke.

11Distance from the equator was measured in terms of degrees latitude, but we ran the analyses with minimum temperature of the day of assessment as well in all our analyses and found comparable results.

Figure 3). The mediation showing the relationship between DEQ, CSI, and CBT survived all analyses for those with a relationship (https://osf.io/jqqeq/), but not so for those without a relationship (https://osf.io/5vypb/). Importantly, having a relationship moderates the link between DEQ and CSI (B = -.020 , t = -4.99 , p < .001, 95% CI [-0.028, -0.012]), but not between CSI and CBT (this moderation survived our exploration dataset, B = -.062 , t = -2.38 , p = .02, 95% CI [-0.113, -0.109] ; but not our test dataset, B = -.027 , t = -1.01, p = .31 , 95% CI [-0.079, 0.025]).

Unpacking this further shows (Figure 5) that for people with a romantic relationship, the association between DEQ and CSI is significant, whereas it is not significant for those without a relationship (Figure 6). For the latter, DEQ relates to lower CBT irrespective of CSI. Conversely, for people with a romantic relationship, the link of DEQ with CBT is to a significant degree mediated by CSI. We thus infer that core body temperature is buffered through complex social integration, and having a romantic relationship seems indeed to indicate an ability to engage in, and extend, the social network to generate warmth. That the mediation is not complete reasonably suggests that other regulation mechanisms also play part in buffering body temperature.

There is yet another way to test for the role of complex social integration, and how humans may be distinct from the vervet monkeys whose core temperature is protected by the *size* of the social network. On the basis of the existing literature (Cohen & Lemay, 2007; Holt-Lunstad et al., 2010) and our machine learning results, we suspected that, with regard to physical effects of social warmth in humans, the *quality* of one's networks (i.e., complex social integration) is superior to the sheer size of people's networks. Of course, CSI correlates considerably with the size of people's social network ($r = .514$, $p < .001$). Yet, if it is indeed complex social integration rather than the sheer size of the network, then network size should not mediate the relationship between DEQ and CBT. Indeed, the mediation did not even survive our first exploratory analyses, as network size was not predicted by DEQ ($B = .09$, $t = 1.10$, $p = .59$, 95% CI [-0.072, 0.257]) nor did network size predict CBT ($B = -.0011$, $t = -.72$, $p = .47$, 95% [-.0043, .0020]). It is thus quality of the network, and not size, that matters.

In short, we infer that maintaining one's core temperature is an important driver for complex social integration and thereby also has consequences for physical, social, and emotional functioning. At least for relational motivations rooted in social thermoregulation, the results suggest that being closer to the equator makes one feel warmer and thus leads to less urgency to engage in complex social integration. Being further away from the equator matters for the degree to which one engages in a higher level of complex social integration.

General Discussion

To our knowledge, the Human Penguin Project (HPP) is the first larger scale study to investigate the interrelation of distance from the equator (DEQ), complex social integration (CSI) and core body temperature (CBT) empirically: In pilot study and a main, cross-national study spanning 12 countries and 3 continents, with various distances form the equator, we find (a) a considerable association between distance from the equator and CSI, and (b) a significant association between CSI and people's CBT for those who are seemingly not inhibited to socially connect (i.e., in our study: those with a romantic relationship). The data are very clear: CSI is closely intertwined with thermoregulation. We infer that many of our "older" systems (like body temperature regulation) are still crucial in shaping our modern ways of connecting with each other. What is also very clear from our data is that DEQ and CBT in relation to CSI are an important part of the story, but not the entire story. Culture (language family) plays a role that may be even more important for CSI, opening up the door to investigate interrelationships between socio-economic development, level of close-knittedness in cultures (Van de Vliert & Lindenberg, 2006), and linguistic structures (Koptjevskaja-Tamm, 2015) with complex social integration and temperature regulation.

We need to be clear: Our studies do not allow for direct causal statements. Theoretically and intuitively however it is unlikely that people's core body temperature drives how far they live from the equator. It is also relatively likely that the level of social integration is predictive of core body temperatures (and not the other way around). We therefore make a relatively cautious inference that Complex Social Integration protects from the cold.

Our results are robust, but the mechanisms via which people arrive at a higher core body temperature in colder environments not yet. Why do complex social networks protect our bodies from the cold? We have answered this question indirectly in our theoretical introduction: People's modern forms of relationships probably grow out of more ancient relationships. That means that people used to huddle with each other, and that in modern relationships we still "track" people's trustworthiness and predictability by gaging whether they are cold or warm. This is confirmed in research showing that lower temperatures increase our desire to more frequently email or call loved ones increases when cold (Van Acker et al., 2016) and research showing that "priming" people with loneliness lead them to estimate ambient temperature as lower (IJzerman & Semin, 2010).

But the more proximate mechanisms are not yet clear. We strongly suspect that direct co-thermoregulatory mechanisms exist. There are some indications that mothers increase their peripheral temperatures when their infants are in distress (Vuorenkoski et al. 1969). In adults, Wagemans and IJzerman (2014) found that people respond with peripheral temperature increases when seeing their sad partner, arguably to co-regulate their partner. The relationship literature is further replete with suggestions that people physiologically co-regulate in the service of homeostasis, which we have argued to include temperature homeostasis (IJzerman, Heine, et al., 2017). From this perspective, temperature regulation has become implicated in attachment processes, which, in turn, form the basis from which people form predictions about others. Vergara et al. (2017) provide the first evidence that social thermoregulatory habits predict individual differences in attachment styles; IJzerman Heine, et al. (2017) outline how to exactly investigate these co-thermoregulatory dynamics.

Conclusion

Although social thermoregulation is a hotly debated topic, the results from the HPP proved to be robust and open up new perspectives to investigate such co-thermoregulatory dynamics. We anticipate our study to be a starting point for other larger scale studies on the connection between temperature regulation, relationship quality, social integration, and health. In order to better understand the role of temperature in relationship (co-)regulation, future studies should assess social thermoregulation itself in more detailed ways, for example through longitudinal studies relying on modern sensor and actuator technologies. For this future work, our HPP study has made a crucial first step towards understanding how human social thermoregulation affects the relationship between complex social integration and our health.

References:

Anderson, M. L. (2010). Neural reuse: A fundamental organizational principle of the brain. *Behavioral and Brain Sciences*, *33*, 245-266.

Benzinger, T. H. (1969). Clinical temperature: New physiological basis. *JAMA, 209,* 1200 - 1206.

Bergmann, C. (1847) Uber die Verhältnisse der Wärmeökonomie der Thiere zu ihrer Grösse. *Gottinger Studien, 3*, 595–708.

Boulant, J. A. (2000). Role of the preoptic-anterior hypothalamus in thermoregulation and fever. *Clinical Infectious Diseases*, *31*, S157-S161.

Bowlby, J. (1969) *Attachment and Loss*. London: Basic Books.

Brandt, M. J. (2017). Predicting ideological prejudice. *Psychological Science*, *28,* Available online at http://journals.sagepub.com/doi/full/10.1177/0956797617693004.

Breiman, L. (1996). Bagging predictors. *MachineLlearning*, *24*, 123-140.

Breiman, L. (2001). Random forests. *Machine Learning, 45,* 5 - 32.

Bylander, T., & Hanzlik, D. (1999) Estimating generalization error using out-of-bag estimates. *AAAI-99 Proceedings.* Available at http://www.aaai.org/Papers/AAAI/1999/AAAI99-047.pdf.

Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, *36*, 181-204.

Cohen, S., & Lemay, E. P. (2007). Why would social networks be linked to affect and health practices? *Health Psychology*, *26*, 410.

Cohen, S., Doyle, W. J., Skoner, D. P., Rabin, B. S., & Gwaltney, J. M., Jr (1997). Social ties and susceptibility to the common cold. *JAMA, 277,* 1940 - 1944.

Comtois, D. (2016). Package ‘summarytools’. Available at https://cran.r-project.org/web/packages/summarytools/summarytools.pdf.

Craig, A. D., Chen, K., Bandy, D., & Reiman, E. M. (2000). Thermosensory activation of insular cortex. *Nature Neuroscience*, *3*, 184-190.

Crocker, J., Luhtanen, R. K., Cooper, M. L., & Bouvrette, A. (2003). Contingencies of self-worth in college students: theory and measurement. *Journal of Personality and Social Psychology*, *85*, 894-908.

Deci, E. L., & Ryan, R. M. (1995). Human autonomy. In *Efficacy, Agency, and Self-esteem* (pp. 31-49). Springer US.

Ebersole, C. R., Atherton, O. E., Belanger, A. L., Skulborstad, H. M., Allen, J. M., Banks, J. B., ... & Brown, E. R. (2016). Many Labs 3: Evaluating participant pool quality across the academic semester via replication. *Journal of Experimental Social Psychology*, *67*, 68-82.

Ebensperger, L. A. (2001). A review of the evolutionary causes of rodent group-living. *Acta Theriologica, 46,* 115 - 144.

Foster, F., & Collard, M. (2013). A reassessment of Bergmann's rule in modern humans. *PloS One*, *8*, e72269.

Fraley, R. C., Waller, N. G., & Brennan, K. A. (2000). An item response theory analysis of self-report measures of adult attachment. *Journal of Personality and Social Psychology*, *78*, 350-365.

Friston, K. (2008). Hierarchical models in the brain. *PLoS Computational Biology*, *4*, e1000211.

Gallup, A. C., & Gallup, G. G. (2008). Yawning and thermoregulation. *Physiology & Behavior*, *95*, 10-16.

Hadi, R., Block, L., & King, D. (2012). Mental thermoregulation: Affective and cognitive pathways for non-physical temperature regulation. *NA-Advances in Consumer Research, 40*. Abstract available online: http://www.acrwebsite.org/volumes/1014152/volumes/v10e/E-10

Harris, P. B., Brown, B. B., & Werner, C. M. (1996). Privacy regulation and place attachment: Predicting attachments to a student family housing facility. *Journal of Environmental Psychology*, *16*, 287-301.

Henriksen, R. E., Torsheim, T., & Thuen, F. (2014). Loneliness, social integration and consumption of sugar-containing beverages: testing the social baseline theory. *PloS one*, *9*, e104421.

Hill, J. R., & Rahimtulla, K. A. (1965). Heat balance and the metabolic rate of newborn babies in relation to environmental temperature and the effect of age and of weight on basal metabolic rate. *The Journal of Physiology, 180,* 239 - 265.

Hlavac, M. (2015). Stargazer. Available at https://cran.r-project.org/web/packages/stargazer/stargazer.pdf.

Hong, J., & Sun, Y. (2012). Warm it up with love: The effect of physical coldness on liking of romance movies. *Journal of Consumer Research*, *39*, 293-306.

Holt-Lunstad, J., Smith, T. B., & Layton, J. B. (2010) Social relationships and mortality risk: A meta-analytic review. *PLoS Med, 7,* e1000316.

Kooiman, C. G., Spinhoven, P., & Trijsburg, R. W. (2002). The assessment of alexithymia: A critical review of the literature and a psychometric study of the Toronto Alexithymia Scale-20. *Journal of Psychosomatic Research*, *53*, 1083-1090.

IJzerman, H., & Hogerzeil, L. J. (2016). People as penguins: Thermoregulation as part of the human essence. In M. van Zomeren & J. Dovidio. *Oxford Handbook of Human Essence*. Oxford University Press. Preprint available at http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2844963.

IJzerman, H., & Semin, G. R. (2009). The thermometer of social relations: Mapping social proximity on temperature. *Psychological Science*, *20*, 1214-1220.

IJzerman, H., & Semin, G. R. (2010). Temperature perceptions as a ground for social proximity. *Journal of Experimental Social Psychology*, *46*, 867-873.

IJzerman, H., Janssen, J. A., & Coan, J. A. (2015). Maintaining warm, trusting relationships with brands: increased temperature perceptions after thinking of communal brands. *PloS One, 10,* e0125194.

IJzerman, H., Szymkow, A., & Parzuchowski, M. (2016). Warmer hearts, and warmer, but noisier rooms: Communality does elicit warmth, but only for those in colder ambient temperatures—Commentary on Ebersole et al.(2016). *Journal of Experimental Social Psychology*, *67*, 88-90.

IJzerman, H., Heine, E. C., Nagel, S. K., & Pronk, T. M. (2017). Modernizing relationship therapy through Social Thermoregulation Theory: Evidence, hypotheses, and explorations. *Frontiers in Psychology*, *8*.

IJzerman, H., Pollet, T., Ebersole, C., & Kun, D. (2016). What predicts Stroop performance? A conditional random forest approach.Available at SSRN: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2805205.

IJzerman, H., Coan, J. A., Wagemans, F. M., Missler, M. A., van Beest, I., Lindenberg, S., & Tops, M. (2014). A theory of social thermoregulation in human primates. *Frontiers in Psychology, 6,* 464.

IJzerman, H., Gallucci, M., Pouw, W. T., Weiβgerber, S. C., Van Doesum, N. J., & Williams, K. D. (2012). Cold-blooded loneliness: Social exclusion leads to lower skin temperatures. *Acta Psychologica, 140,* 283 – 288.

Inagaki, T. K., & Eisenberger, N. I. (2013). Shared neural mechanisms underlying social warmth and physical warmth. *Psychological Science*, *24,* 2272 – 2280.

Inagaki, T. K., Irwin, M. R., & Eisenberger, N. I. (2015). Blocking opioids attenuates physical warmth-induced feelings of social connection. *Emotion*, *15*, 494.

Inagaki, T. K., Irwin, M. R., Moieni, M., Jevtic, I., & Eisenberger, N. I. (2016). A pilot study examining physical and social warmth: higher (non-febrile) oral temperature is associated with greater feelings of social connection. *PloS One*, *11*, e0156873.

Jansky, L. (1973). Non-shivering thermogenesis and its thermoregulatory significance. *Biological Reviews, 48,* 85-132.

Jones, Z., & Linder, F. (2015). Exploratory data analysis using random forests. *Prepared for the 73rd annual MPSA conference*. Available online at http://zmjones.com/static/papers/rfss_manuscript.pdf.

Kernis, M. H. (2003). Toward a conceptualization of optimal self-esteem. *Psychological Inquiry*, *14*, 1-26.

Koptjevskaja-Tamm, M. (2015). *The linguistics of temperature.* Amsterdam: John Benjamins Publishing Company).

LeBel, E. P., & Campbell, L. (2013). Heightened sensitivity to temperature cues in individuals with high anxious attachment. Real or elusive phenomenon?. *Psychological Science*, *24,* 2128 – 2130.

Marazziti, D., Di Muro, A., & Castrogiovanni, P. (1992). Psychological stress and body temperature changes in humans. *Physiology & Behavior*, *52*, 393-395.

McFarland, R., Fuller, A., Hetem, R. S., Mitchell, D., Maloney, S. K., Henzi, S. P., & Barrett, L. (2015). Social integration confers thermal benefits in a gregarious primate. *Journal of Animal Ecology, 84,* 871 - 878.

McFarland, R., Henzi, S. P., Barrett, L., Wanigaratne, A., Coetzee, E., Fuller, A., ... & Maloney, S. K. (2015). Thermal consequences of increased pelt loft infer an additional utilitarian function for grooming. *American Journal of Primatology, 78,* 456 - 461.

Minard, D., Copman, L., & Dasler, A. R. (1964). Elevation of body temperature in health. *Annals of the New York Academy of Sciences, 121,* 12 - 25.

Mizukami, K., Kobayashi, N., Ishii, T., & Iwata, H. (1990). First selective attachment begins in early infancy: A study using telethermography. *Infant Behavior and Development,* 13, 257 - 271.

Nuñez-Villegas, M., Bozinovic, F., & Sabat, P. (2014). Interplay between group size, huddling behavior and basal metabolism: an experimental approach in the social degu. *Journal of Experimental Biology, 217,* 997 - 1002.

Open Science Collaboration. (2015). Estimating the reproducibility of psychological science. *Science*, *349*, aac4716.

Peters, R. H. (1986). *The ecological implications of body size*. Cambridge University Press.

Ripley, B. (2016). Package 'tree'. Available at https://cran.r-project.org/web/packages/tree/tree.pdf.

Rotman, J. D., Lee, S. H. M., & Perkins, A. W. (2016). The warmth of our regrets: Managing regret through physiological regulation and consumption. *Journal of Consumer Psychology*, *2,* 160 – 170.

Routledge, C., Arndt, J., Sedikides, C., & Wildschut, T. (2008). A blast from the past: The terror management function of nostalgia. *Journal of Experimental Social Psychology*, *44*, 132-140.

Sarkar, D. (2017). Lattice: Trellis Graphics for R. Available at https://cran.r-project.org/web/packages/lattice/index.html.

Satinoff, E. (1983). A reevaluation of the concept of the homeostatic organization of temperature regulation. In *Motivation* (pp. 443-472). Springer US.

Satinoff, E. (1982). Are there similarities between thermoregulation and sexual behavior?. In *The physiological mechanisms of motivation* (pp. 217-251). Springer New York.

Schilder, J. D., IJzerman, H., & Denissen, J. J. (2009). Physical warmth and perceptual focus: A replication of IJzerman and Semin (2009). *PloS One, 9,* e112772.

Seeman, T. E. (1996). Social ties and health: The benefits of social integration. *Annals of Epidemiology*, *6*, 442-451.

Steverink, N. & Lindenberg, S. (2008). Do good self-managers have less physical and social deficits and more well-being in later life? *European Journal of Ageing,* 5: 181-190

Strobl, C., Malley, J., & Tutz, G. (2009). An introduction to recursive partitioning: Rationale, application, and characteristics of classification and regression trees, bagging, and random forests. *Psychological Methods*, *14*, 323-348.

Tangney, J. P., Baumeister, R. F., & Boone, A. L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality, 72,* 271 – 324.

Terrien, J., Perret, M., & Aujard, F. (2011). Behavioral thermoregulation in mammals: a review. *Frontiers in Bioscience*, *16*, 1428-1444.

Umberson, D. (1987) Family status and health behaviors: Social control as a dimension of social integration. *Journal of Health and Social Behavior, 28*, 306 - 319.

Van Acker, B., Kerselaers, K., Pantophlet, J., & IJzerman, H. (2016). Homelike thermoregulation: How physical coldness makes an advertised house a home. *Journal of Experimental Social Psychology, 67,* 20 – 27.

Van de Vliert, E., & Lindenberg, S, (2006). Wealth, climate, and framing: Cross-national differences in solidarity. In D. Fetchenhauer, A. Flache, A. P. Buunk, & S. Lindenberg (eds.). *Solidarity and Prosocial Behavior: An Integration of Sociological and Psychological Perspectives.* New York: Springer.

Vergara, R. C., Hernández, C., Jaume-Guazzini, F., Lindenberg, S., & IJzerman, H. (2017). Development and validation of the Social Thermoregulation and Risk Avoidance Questionnaire (STRAQ-1). Unpublished dataset.

Vess, M. (2012). Warm thoughts: Attachment anxiety and sensitivity to temperature cues. *Psychological science*, *23*, 472-474.

Vuorenkoski, V., Wasz-Hockert, O., Koivisto, E., & Lind, J. (1969). The effect of cry stimulus on the temperature of the lactating breast of primipara. A thermographic study. *Experientia, 25,* 1286–1287.

Wagemans, F. A. M., & IJzerman, H. (2014). Social thermoregulation in close relationships. Poster presented at SPSP; available at Figshare https://figshare.com/articles/Social_thermoregulation_in_close_relationships/926484.

Wickham, H. (2016). Package ‘plyr’. Available at https://cran.r-project.org/web/packages/plyr/index.html.

Williams, L E, & Bargh, J A (2008). Experiencing physical warmth promotes interpersonal warmth. *Science, 322,* 606 - 607.

Yarkoni, T., & Westfall, J. (in press). Choosing prediction over explanation in psychology: Lessons from machine learning. *Perspective on Psychological Science. Preprint available at https://dx. doi. org/10.6084/m9.*

Zhou, X., Wildschut, T., Sedikides, C., Chen, X., & Vingerhoets, A. J. (2012). Heartwarming memories: Nostalgia maintains physiological comfort. *Emotion, 12*, 678-684.

Zwebner, Y., Lee, L., & Goldenberg, J. (2013). The temperature premium: Warm temperatures increase product valuation. *Journal of Consumer Psychology*, *24*, 251-259.

Figures

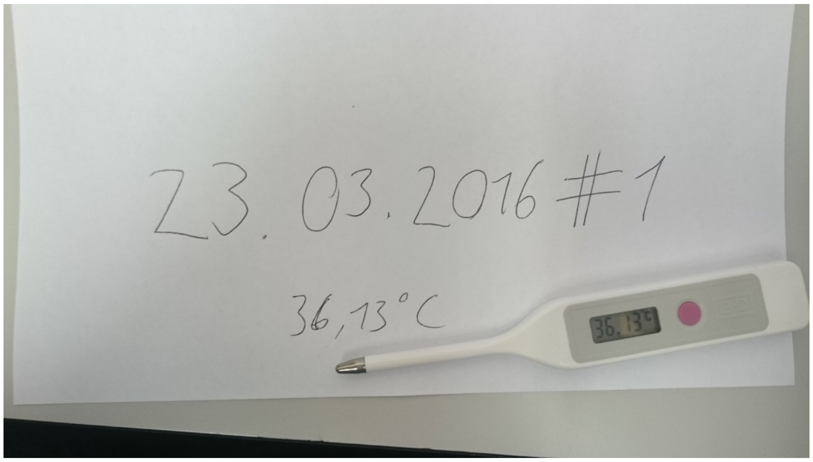

Figure 1 - HPP Thermometer. *An example picture of a thermometer uploaded to Qualtrics.*

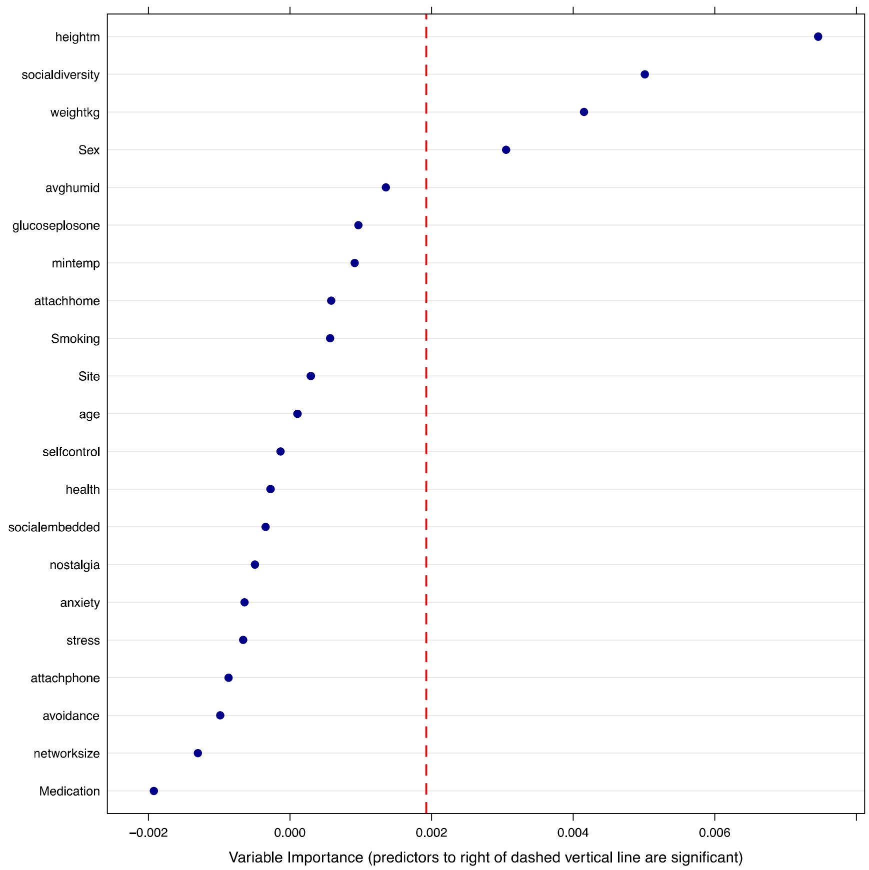

Figure 2 – HPP Pilot Dotplot. *Permutation variable importance of predictors of Core Body Temperature from our supervised machine learning analyses in our pilot study. Variables exceeding the red line are very unlikely random noise.*

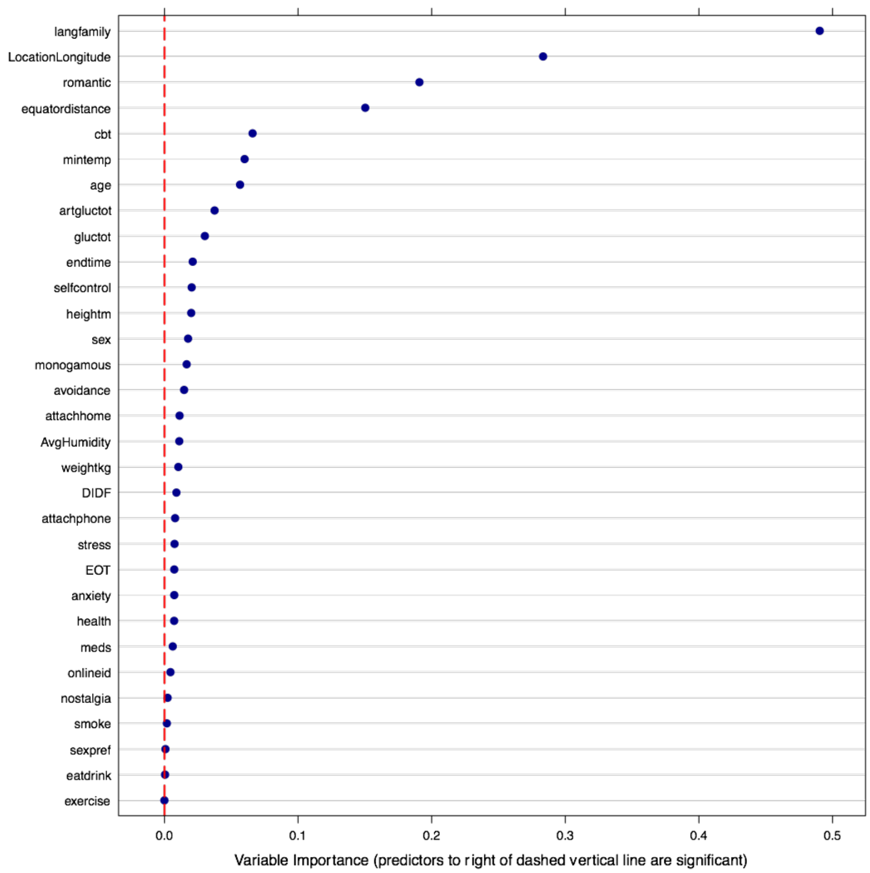

Figure 3 HPP Cross-National Study Dotplot for CSI. *Permutation variable importance of predictors of Complex Social Integration from our supervised machine learning analyses in our pilot study. Variables exceeding the red line are very unlikely random noise.*

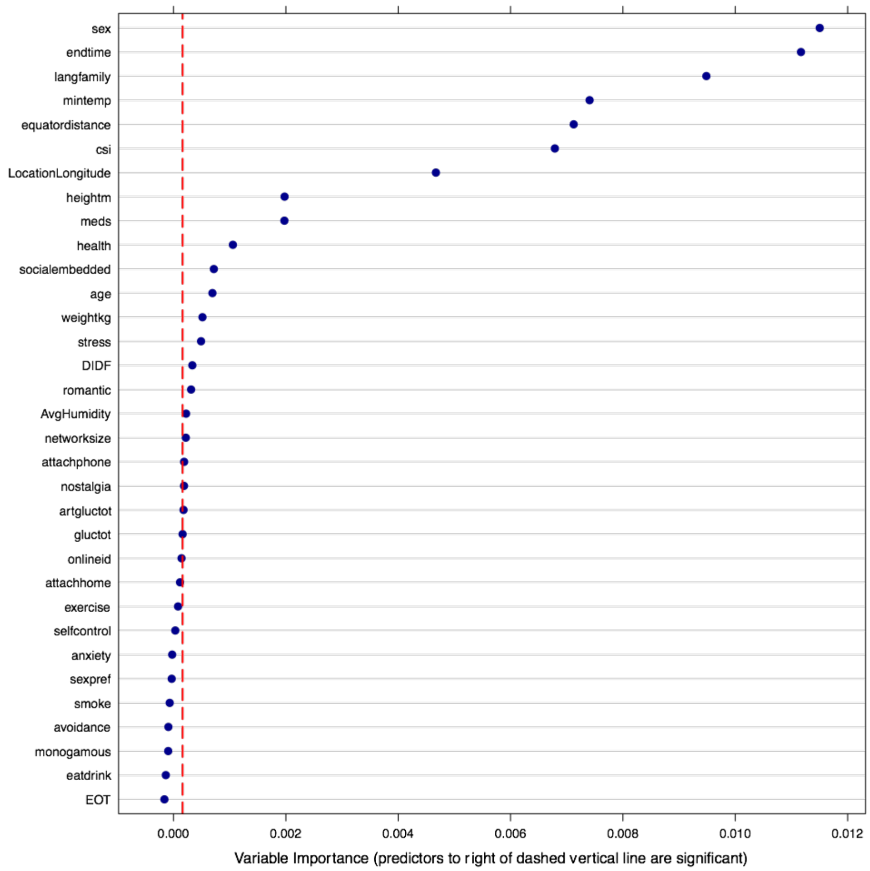

Figure 4 HPP Cross-National Study Dotplot for CBT. *Permutation variable importance of predictors of Core Body Temperature from our supervised machine learning analyses in our pilot study. Variables exceeding the red line very likely differ from random noise.*

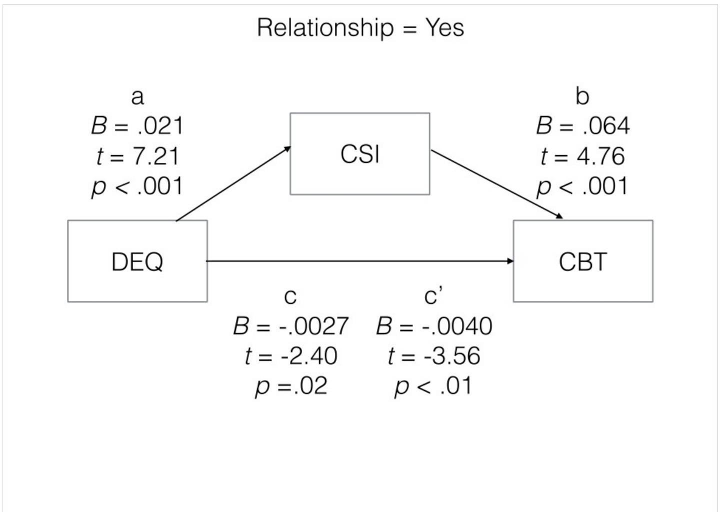

Figure 5 HPP Mediation Model for people in a relationship. *Mediation analyses showing how Complex Social Integration (CSI) protects the core temperatures (CBT) of people with a romantic relationship from colder climates (DEQ), c = significance with, and c' = significance without mediation by CSI.*

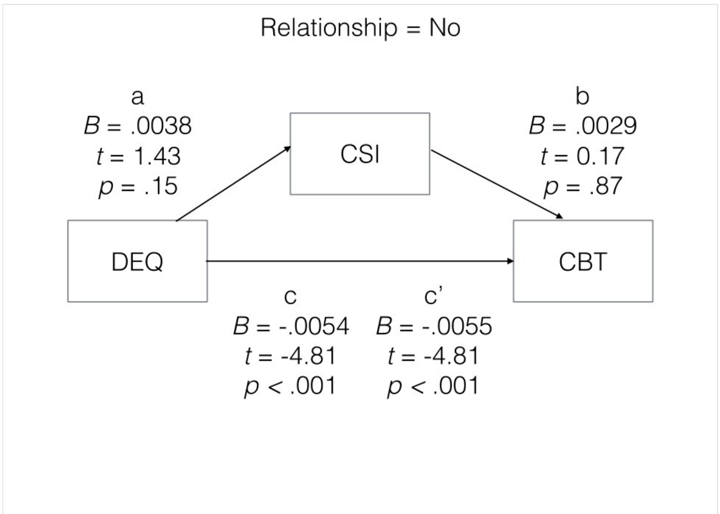

Figure 6 HPP Mediation Model for people not in a relationship. *Mediation analyses showing that Complex Social Integration (CSI) does not protect the core temperatures (CBT) of people without a romantic relationship from colder climates (DEQ), c = significance with, and c’ = significance without mediation by CSI.*